

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: 06309362 A

(43) Date of publication of application: 04 . 11 . 94

(51) Int. Cl

G06F 15/40  
G06F 15/40

(21) Application number: 05101348

(22) Date of filing: 27 . 04 . 93

(71) Applicant: FUJITSU LTD

(72) Inventor: KIKUTA YASUYO  
TAZAWA KAZUHIKO  
MURATA AKIRA  
OTA TATSUO  
AKAMATSU YURI

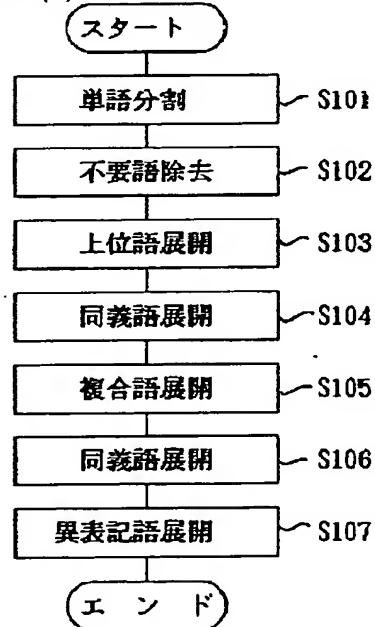
(54) INFORMATION RETRIEVING METHOD

(57) Abstract:

PURPOSE: To shorten the retrieval time while obtaining a high hit rate at the time of retrieval as to the information retrieving means which retrieves information from a book data base for an information processing field by using a key word for retrieval.

CONSTITUTION: When book data are registered, the book data to be registered are divided (101) into words by using a morpheme dictionary, a word which is not necessary as a key word is removed by using an unnecessary word dictionary to employ the result as the key word (102), and a high-order word for the key word is developed (103) by using a high-order word dictionary; and a synonym for the key word which is already generated is developed (104) by using a synonym dictionary, a compound word for the key word which is already generated is divided (105) by using a compound dictionary, and a synonym for the key word generated by dividing the compound word is developed (106) by using the synonym dictionary. Then a new key word is generated (107) by developing a different-notation word for the key word which is already generated by using a different-notation word dictionary.

COPYRIGHT: (C)1994,JPO



(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平6-309362

(43)公開日 平成6年(1994)11月4日

(51)Int.Cl.<sup>5</sup>  
G 0 6 F 15/40

識別記号 庁内整理番号  
5 0 0 Q 9194-5L  
5 1 0 P 9194-5L

F I

技術表示箇所

審査請求 未請求 請求項の数 4 OL (全 13 頁)

(21)出願番号 特願平5-101348

(22)出願日 平成5年(1993)4月27日

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中1015番地

(72)発明者 菊田 泰代

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(72)発明者 田澤 和彦

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(72)発明者 村田 彰

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(74)代理人 弁理士 伊東 忠彦

最終頁に続く

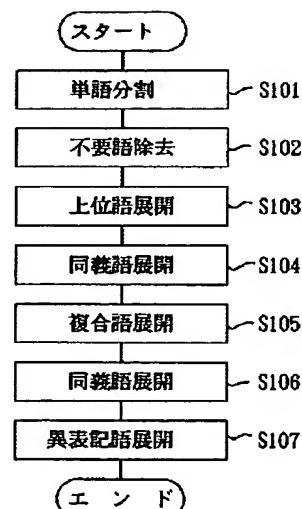
(54)【発明の名称】 情報検索方法

(57)【要約】

【目的】 本発明は、検索用キーワードを用いて情報処理分野の文献データベースより情報を検索する情報検索方法に関し、検索時の高いヒット率を実現しながら、検索時間を短縮することを目的とする。

【構成】 文献データの登録時に、形態素辞書を用いて登録する文献データを単語に分割し(101)、不要語辞書を用いてキーワードとして不要な単語を除去してキーワードとして採用し(102)、上位語辞書を用いて上記キーワードに対する上位語を展開し(103)、同義語辞書を用いて生成済キーワードに対して同義語を展開し(104)、複合語辞書を用いて生成済キーワードに対して複合語を分割し(105)、同義語辞書を用いて複合語の分割で生成されたキーワードに対して同義語を展開し(106)、異表記語辞書を用いて生成済キーワードに対して異表記語を展開して新たなキーワードを生成する(107)。

第1実施例におけるキーワードの  
展開手順を示すフローチャート



**【特許請求の範囲】**

**【請求項1】** 形態素辞書(15)を用いて、文献データベースに登録する文献データの文章を単語に分割するステップ(101)と、  
 不要語辞書(16)を用いて、上記分割された単語から、検索用キーワードとして不要な単語を除去して、残った単語を検索用キーワードとして採用するステップ(102)と、  
 上位語辞書(18)を用いて、上記採用されたキーワードに対する上位語を展開して新たな検索用キーワードを生成するステップ(103)と、  
 同義語辞書(19)を用いて、生成済の検索用キーワードに対して、読みは異なるが同じ意味の同義語を展開して新たな検索用キーワードを生成するステップ(104)と、  
 複合語辞書(20)を用いて、生成済の検索用キーワードに対して、複合語を分割して新たな検索用キーワードを生成するステップ(105)と、  
 上記同義語辞書(19)を用いて、上記複合語の分割により生成された検索用キーワードに対して、読みは異なるが同じ意味の同義語を展開して新たな検索用キーワードを生成するステップ(106)と、  
 異表記語辞書(21)を用いて、生成済の検索用キーワードに対して、本来同じ単語であるが表記のみが異なる異表記語を展開して新たな検索用キーワードを生成するステップ(107)とを有し、  
 上記文献データベースに、上記文献データと共に上記生成した検索用キーワードを登録しておき、文献データの検索時には、検索者により入力された検索用キーワードを用いて、目的とする文献データを検索する構成としたことを特徴とする情報検索方法。

**【請求項2】** 前記上位語辞書(18)、同義語辞書(19)、複合語辞書(20)、及び異表記語辞書(21)は、検索用キーワードとして有用なものだけが登録されていることを特徴とする請求項1記載の情報検索方法。

**【請求項3】** 文献データの検索時に、入力された検索用キーワードにより文献データを検索した結果、検索用キーワードに該当する文献データの件数が所定値以上であるかどうかを判断するステップ(203)と、  
 上記判断の結果、検索用キーワードに該当する文献データの件数が基準値以上である場合には、下位語展開により下位概念の検索用キーワードの候補を生成し、上記検索用キーワードの候補を検索者に提示して、検索者に検索用キーワードを選択させるステップ(204、205)を有することを特徴とする請求項1記載の情報検索方法。

**【請求項4】** 文献データの検索時に検索者により入力された自然語の質問文を、前記形態素辞書(15)により単語に分割し、前記不要語辞書(16)により不要な

単語を辞去して検索用キーワードとして有用な単語のみを残して、検索用キーワードの候補とするステップ(302、303、304)と、  
 検索者に、上記検索用キーワードの候補を提示して確認させ、必要に応じて検索用キーワードを変更させて検索用キーワードを決定するステップ(305)とを有し、上記決定された検索用キーワードを用いて文献データを検索することを特徴とする請求項1記載の情報検索方法。

**【発明の詳細な説明】****【0001】**

**【産業上の利用分野】** 本発明は情報検索方法に係り、特に、検索用キーワードを用いて情報処理分野の文献データベースより情報を検索する情報検索方法に関する。

**【0002】** 文献データベースより情報を検索する情報検索方法では、目的とする文献データのヒット率の向上、検索時間の短縮が必要とされている。

**【0003】**

**【従来の技術】** 文献データベースの検索を行う情報検索

装置では、文献データの登録時に、登録する文献データにキーワードを付与して情報検索装置に登録しておき、検索時には、キーワードを入力して、目的の文献データを検索している。

**【0004】** 従来、検索時のヒット率を向上させるために、検索時に、同義語辞書、上位語辞書、下位語辞書等の展開語辞書を用いて、検索者により入力された元のキーワードを展開して関連するキーワードを生成し、この生成した関連するキーワードにより検索を行う方法がある。

**【0005】**

**【発明が解決しようとする課題】** しかし、従来の、検索時に展開語辞書を用いてキーワードを展開して、生成した関連するキーワードを用いて検索を行う方法では、キーワードの展開に時間がかかるため、検索に要する時間が長いという問題がある。

**【0006】** また、文献データの登録時には、人手でキーワードを付けるため、ヒット率の向上を可能とする適切なキーワードを付けることが難しく、手間もかかるという問題がある。

**【0007】** 本発明は、上記の点に鑑みてなされたもので、検索時の高いヒット率を実現しながら、検索時間を短縮することができる情報検索方法を提供することを目的とする。

**【0008】**

**【課題を解決するための手段】** 請求項1の発明は、形態素辞書を用いて、文献データベースに登録する文献データの文章を単語に分割するステップと、不要語辞書を用いて、上記分割された単語から、検索用キーワードとして不要な単語を除去して、残った単語を検索用キーワードとして採用するステップと、上位語辞書を用いて、上

記採用されたキーワードに対する上位語を展開して新たな検索用キーワードを生成するステップと、同義語辞書を用いて、生成済の検索用キーワードに対して、読みは異なるが同じ意味の同義語を展開して新たな検索用キーワードを生成するステップと、複合語辞書を用いて、生成済の検索用キーワードに対して、複合語を分割して新たな検索用キーワードを生成するステップと、上記同義語辞書を用いて、上記複合語の分割により生成された検索用キーワードに対して、読みは異なるが同じ意味の同義語を展開して新たな検索用キーワードを生成するステップと、異表記語辞書を用いて、生成済の検索用キーワードに対して、本来同じ単語であるが表記のみが異なる異表記語を展開して新たな検索用キーワードを生成するステップとを有し、上記文献データベースに、上記文献データと共に上記生成した検索用キーワードを登録しておき、文献データの検索時には、検索者により入力された検索用キーワードを用いて、目的とする文献データを検索する構成とする。

【0009】請求項3の発明は、文献データの検索時に、入力された検索用キーワードにより文献データを検索した結果、検索用キーワードに該当する文献データの件数が所定値以上であるかどうかを判断するステップと、上記判断の結果、検索用キーワードに該当する文献データの件数が基準値以上である場合には、下位語展開により下位概念の検索用キーワードの候補を生成し、上記検索用キーワードの候補を検索者に提示して、検索者に検索用キーワードを選択させるステップを有する構成とする。

【0010】請求項4の発明は、文献データの検索時に検索者により入力された自然語の質問文を、前記形態素辞書により単語に分割し、前記不要語辞書により不要な単語を削除して検索用キーワードとして有用な単語のみを残して、検索用キーワードの候補とするステップと、検索者に、上記検索用キーワードの候補を提示して確認させ、必要に応じて検索用キーワードを変更させて検索用キーワードを決定するステップとを有し、上記決定された検索用キーワードを用いて文献データを検索する構成とする。

#### 【0011】

【作用】請求項1の発明では、文献データの登録時に、展開語辞書により自動的に適切な検索用キーワードを生成して登録する。このため、検索時の高いヒット率を実現しながら、検索時間を短縮することを可能とし、また、検索用キーワードを付与する手間を削減することを可能とする。

【0012】請求項3の発明では、検索用キーワードによる文献データの検索の結果、該当する文献データの件数が基準値以上の場合には、下位概念の検索用キーワード候補を生成して検索者に選択させて、検索対象を絞り込む。このため、検索用キーワードに該当する文献データ

タの件数が多い場合に、目的の文献データを検索するまでの検索時間を短縮することを可能とする。

【0013】請求項4の発明では、検索時に、検索者により入力された自然語の質問文から、適切なキーワードを生成する。このため、容易に適切なキーワードを生成させることができ、検索時のヒット率を向上させることを可能とする。

#### 【0014】

【実施例】図1は本発明の第1実施例を適用した情報検索装置の構成図を示す。この情報検索装置は、情報処理分野の文献データベースを登録し、検索する装置の例である。図1に示す情報検索装置は、登録する文献データを入力する文献データ入力装置12、検索者が検索用キーワードやコマンド等を入力する入力装置13、検索結果等を表示する表示装置14、文献データとキーワードの登録処理、検索処理を行う処理部11を備えている。

【0015】また、補助記憶装置等に格納されるデータベース25は、複数の文献データからなる文献データファイル26と、各文献データに付与された検索用キーワードの集まりでありキーワードファイル27からなる。キーワードファイル27中の各キーワードは、対応する文献データの格納位置を示すポインタと共に記憶されている。

【0016】また、文献データ登録時に、文献データの文章からキーワードを抽出して、抽出した各キーワードを基に、関連する単語にキーワードを展開して検索用キーワードを生成するために、形態素辞書15、不要語辞書16、展開語辞書17を備えている。展開語辞書17は、上位語辞書18、同義語辞書19、複合語辞書20、異表記語辞書21からなる。上記の各辞書は、補助記憶装置等に格納されている。

【0017】形態素辞書15には、各種文字列と品詞(名詞、接続詞、助詞等)の対応を示す情報が記憶されている。不要語辞書16には、接続詞、助詞等の不要な単語、及び、情報処理分野でキーワードとして無用な単語が登録されている。例えば、“ソフト”という単語は、極一般的な単語であるため、不要語として登録されている。

【0018】また、上位語辞書18には、情報処理分野で使われる各種の単語について、上位概念の単語、下位概念の単語を示す情報が記憶されている。同義語辞書19には、情報処理分野で使われる各種の単語について、読みは異なるが同じ意味の同義語を示す情報が記憶されている。

【0019】複合語辞書20には、情報処理分野で使われる各種複合語について、複合語を分割した単語が登録されている。

【0020】異表記語辞書21には、情報処理分野で使われる各種単語について、本来同じ単語であるが表記のみが異なる異表記語を示す情報が記憶されている。互い

に異表記語の関係にあるものとしては、例えば、一つの単語を、アルファベットで表記した単語と、カタカナで表記した単語がある。

【0021】なお、上記展開語辞書17には、キーワードとして無用なものは、登録していない。

【0022】次に、第1実施例における、文献データを登録する際のキーワードの展開について説明する。本実施例では、文献データを登録する際に、処理部11が、文献データの文章から自動的にキーワードを抽出して、抽出した各キーワードを基に、関連する単語にキーワードを展開して検索用キーワードを生成する。

【0023】図2は、第1実施例におけるキーワードの展開手順を示すフローチャートである。また、図3は、第1実施例におけるキーワードの展開の一例の説明図を示す。

【0024】先ず、文献データ入力装置12により入力された文献データの文章を、形態素辞書15を用いて単語に分割する(ステップ101)。次に、不要語辞書16を用いて、接続詞、助詞等の不要な単語、及び、情報処理分野におけるキーワードとして無用な単語を辞去する。この不要語の除去が終了した時点で残った単語を、文献データの文章から抽出されたキーワードとして採用する(ステップ102)。

【0025】この後に続くキーワード展開処理では、ステップ102で抽出された各キーワードに対して、順次、各展開処理を行って、新たなキーワードを生成する。先ず、上位語辞書18を用いて、元のキーワードの上位語を展開して新たなキーワードを生成する。図3の例では、上位語展開113に示すように、元のキーワードである、“LOTUS 1-2-3”に対して、上位語である“スプレッドシート”を新たなキーワードとして生成している(ステップ103)。

【0026】次に、同義語辞書19を用いて、ステップ103までに生成済のキーワードに対して同義語を展開して、新たなキーワードを生成する。図3の例では、同義語展開114に示すように、“スプレッドシート”に対する同義語である、“表計算ソフト”を新たなキーワードとして生成している(ステップ104)。

【0027】次に、複合語辞書20を用いて、ステップ104までに生成済のキーワードに対して複合語を分割して、新たなキーワードを生成する。図3の例では、複合語展開115に示すように、複合語である“表計算ソフト”に対して、分割した“表計算”と、“計算ソフト”を新たなキーワードとして生成している。なお、“表計算ソフト”を分割すると“表”、“計算”、“ソフト”も生じるが、これらは、情報処理分野のキーワードとして無用であるので、複合語辞書20には、登録していない。このように、複合語を分割して生成される単語のうち、キーワードとして無用なものを登録しないでおくことにより、余分なキーワードを生成することを防

ぐことができ、キーワードファイル27の容量の削減もできる(ステップ105)。

【0028】次に、再度、同義語辞書19を用いて、ステップ105で複合語の分割により生成されたキーワードに対して同義語を展開して、新たなキーワードを生成する。図3の例では、同義語展開116に示すように、“表計算”に対する同義語である、“カルク”を新たなキーワードとして生成している(ステップ106)。

【0029】次に、異表記語辞書21を用いて、ステップ106までに生成済のキーワードに対して異表記語を展開して、新たなキーワードを生成する。図3の例では、異表記語展開117に示すように、“LOTUS 1-2-3”に対する異表記語の“ロータス1-2-3”、及び“スプレッドシート”に対する異表記語の“SPREADSHEET”を、新たなキーワードとして生成している(ステップ107)。

【0030】上記の展開語辞書17を用いたキーワード展開で生成されたキーワードは、登録する文献データのキーワードとして、キーワードファイル27に登録される。一方、文献データは、データベース25の文献データファイル26に登録される。

【0031】上記のように、本実施例では、文献データの登録時に、展開語辞書17を用いて、文献データから抽出した元のキーワードと関連する有用なキーワードを自動的に生成して登録する。このため、文献データの検索時には、キーワードの展開処理をすることなく、広い範囲のキーワードで目的の文献を検索することができる。

【0032】従って、本実施例では、文献データの検索時に、従来方法に比べて、キーワード展開処理に要する時間分、検索時間を短縮することができ、かつ、良好なヒット率を実現することができる。また、人手により検索用キーワードを付与する手間を削減することができる。

【0033】次に、第1実施例における、文献データの検索について説明する。図4は、第1実施例における検索の手順を示すフローチャートである。

【0034】キーワードが検索者により入力されると(ステップ201)、入力されたキーワードを用いて、データベース25より文献データの検索を行う(ステップ202)。次に、検索の結果、キーワードに該当する文献データの件数が基準値以上であるかどうかを判断する(ステップ203)。

【0035】検索結果の件数が基準値以上である場合には、件数を絞るために、下記の処理を行う。先ず、上位語辞書18を逆引きして、元のキーワードの下位語を展開して、下位概念のキーワードを生成する(ステップ204)。次に、生成した下位概念のキーワードの一覧を表示装置14により表示して、検索者に使用する下位概念のキーワードを選択させる(ステップ205)。

【0036】次に、ステップ202に戻って、ステップ205で選択された、元のキーワードの下位概念のキーワードを用いて、再度文献データの検索を行う。以後、検索結果の件数が基準値未満になるまで、上記の処理を続ける。

【0037】ステップ203で、検索結果の件数が基準値未満である場合は、検索結果の文献データの一覧を表示装置14に表示して、検索者に文献データを選択させる（ステップ206）。検索者により文献データが選択されると、この選択された文献データを表示装置14等に出力する（ステップ207）。

【0038】図5は、第1実施例の検索時における下位概念のキーワードの生成例を示す。図5のA欄では、元のキーワードの“製造”に該当する件数が基準値以上で、下位概念のキーワードである、“化学”，“機械”～が生成された例を示している。

【0039】また、図5のB欄では、元のキーワードの“保険”に該当する件数が基準値以上で、下位概念のキーワードである、“生命保険”，“損害保険”～が生成された例を示している。また、図5のC欄では、元のキーワードの“パソコン”に該当する件数が基準値以上で、下位概念のキーワードである、“FMRシリーズ”，“FM-TOWNSシリーズ”～が生成された例を示している。

【0040】上記のように、第1実施例の検索方法では、検索者により入力されたキーワードによる文献データの検索の結果、該当する文献データの件数が基準値以上の場合には、自動的に、下位語展開により下位概念のキーワードを生成して検索者に提示し、検索者に下位概念のキーワードを選択させて、検索対象を絞り込むことができる。このため、キーワードに該当する文献データの件数が多い場合に、目的の文献データを検索するまでの検索時間を短縮することができる。

【0041】なお、検索時にも、文献データ登録時と同様に、キーワードの展開を行う方法とすることも可能である。

【0042】図6は本発明の第2実施例を適用した情報検索装置の構成図を示す。図6において、図1と同一構成部分には同一符号を付し、適宜説明を省略する。この情報検索装置は、情報処理分野の文献データベースを登録し、検索する装置の例である。

【0043】補助記憶装置等には、データベース群24として、複数のデータベース25,～25<sub>n</sub>が格納されている。また、格納されている複数のデータベース25,～25<sub>n</sub>に関する、データベース定義情報22が格納されている。また、検索者により入力される質問文から抽出したキーワードを分類するために、知識ベース23を備えている。また、文献データとキーワードの登録処理、検索処理は、処理部11により行われる。

【0044】図6の情報検索装置では、図1の装置と同

様に、形態素辞書15、不要語辞書16、展開語辞書17を備えており、文献データ登録時には、文献データの文章からキーワードを抽出して、抽出した各キーワードを基に、関連する単語にキーワードを展開して検索用キーワードを生成して登録する。

【0045】第2実施例では、検索時に、検索者に自然語の質問文を入力させ、入力された質問文からキーワードの候補を抽出して検索者に提示し、検索者がこのキーワードを確認した後、文献データの検索を行う。

【0046】図7は、第2実施例における、検索の手順を示すフローチャートである。また、図8、図9は、自然語の質問文の例と、抽出されたキーワードに基づく検索条件の確認時及び検索条件変更時の表示例を示す。

【0047】自然語の質問文が検索者により入力されると（ステップ301）、形態素辞書15を用いて、質問文を単語に分割し（ステップ302）、不要語辞書16を用いて、キーワードとして不要な単語を辞去する（ステップ303）。

【0048】次に、情報処理分野に関する知識ベース20を用いて、不要語を除去して残った単語を、機種、OS（オペレーティングシステム）、業種、業務等の検索キーにより分類して、上記各検索キーに割り振る。この各検索キーに割り振られた単語を、キーワードの候補であるキーワードの選択肢とする。また、各検索キーのキーワードの選択肢のうち、第1候補のキーワードを選定する。

【0049】また、データベース定義情報22を用いて、上記各検索キーの第1候補のキーワードと関係のあるデータベースの種類を調べて、データベースの選択肢を生成する（ステップ304）。

【0050】次に、各検索キーのキーワードの選択肢とデータベースの選択肢を表示装置14に表示して、検索者に選択肢の選択をさせる。この場合、選択肢のうちの第1候補に印をつけて表示する。検索者が、キーワードの選択肢とデータベースの選択肢を選択した時点で、検索条件が決定される（ステップ305）。

【0051】図8の例では、検索者により図8（A）に示す、自然語の質問文、「食品販売店に食品を卸す食品卸業の、物流管理と在庫管理の事例と手法を探したい」が、入力されている。これに対して、図8（B）に示すように、各検索キーごとのキーワードの選択肢と、データベースの選択肢を表示装置14に表示している。

【0052】検索キーの業種には、キーワードの選択肢として、「食品卸」と「食品販売」が表示され、検索キーの業務には、キーワードの選択肢として、「物流管理」と「在庫管理」が表示されている。また、他の検索キーには、キーワードの選択肢として、「事例」と「手法」が表示されている。図8において、四角□は、選択用の枠を示し、黒色の四角■は、現在選択されていることを示す。

【0053】上記キーワードの選択肢を表示する時点では、処理部11により、第1候補として、「食品卸」、「物流管理」、「事例」が選択されている。

【0054】また、データベースの選択肢には、上記各検索キーで現在選択されている選択肢に関するデータベースの種類が表示される。図8(B)では、「事例データベース」、「提案書事例データベース」、「開発事例データベース」～が表示されている。また、第1候補として、処理部11により、「事例データベース」が選択されている。

【0055】検索者は、上記例のように表示されたキーワードの選択肢とデータベースの選択肢を選択して、検索条件を決定する。

【0056】検索者がキーワードの選択肢とデータベースの選択肢を選択して、検索条件が決定されると、処理部11は、文献データの検索を行う。即ち、データベース定義情報22を基に、選択されたデータベース調べて、決定された検索条件に従って、文献データの検索を行う(ステップ306)。この検索の結果、各検索キーのキーワードの組合せに該当する文献データの件数を表示する(ステップ307)。

【0057】次に、検索者に、検索範囲の拡大又は絞り込みのために、検索条件の変更をするかどうかの確認をさせる。検索の結果、該当件数が少なく、検索者が検索条件の変更を選択し、かつ、検索範囲の拡大を選択した場合には、展開語辞書17を用いてキーワードの上位語を展開して、展開した上位概念のキーワードを検索者に選択させる。

【0058】また、検索の結果、該当件数が多く、検索者が検索条件の変更を選択し、かつ、検索範囲の絞り込みを選択した場合には、図4の検索手順で示したように、展開語辞書17を用いてキーワードの下位語を展開して、展開した下位概念のキーワードを検索者に選択させる。

【0059】検索条件が変更されて、検索者により新たなキーワードが選択された場合には、ステップ306に戻って、再度、文献データの検索を行う。

【0060】ステップ308で、検索者が検索条件の変更を選択しなかった場合には、検索結果の文献データの一覧を表示装置14に表示して、検索者に文献データを選択させる(ステップ310)。検索者により文献データが選択されると、この選択された文献データを表示装置14等に出力する(ステップ311)。

【0061】図8の例では、図8(B)に示す検索条件で検索を行った結果、図8(C)に示すように、検索結果の表示がされている。この場合、該当件数が1件で少ないため、図8(C)に示すように、検索者は、検索条件の変更を選択して、検索範囲の拡大を選択し、範囲を広げたい検索キーとして、業種を選択している。

【0062】これにより、処理部11は、業種の元のキ

ーワードである「食品卸」を上位語に展開して、上位概念のキーワードの一覧である、「流通」、「販売」～を表示している。検索者は、この表示された上位概念のキーワードの一覧から、キーワードを選択する。これにより、検索対象を拡大して、再度、検索を行うことができる。

【0063】また、図9の例では検索者により図9(A)に示す、自然語の質問文、「UNIXに関するマニュアルとアプリケーションガイドを探したい」が、入力されている。これに対して、図9(B)に示すように、各検索キーごとのキーワードの選択肢と、データベースの選択肢を表示している。

【0064】検索キーのOSには、キーワードの選択肢として、「UNIX」だけが表示され、他の検索キーには、キーワードの選択肢として、「マニュアル」と「アプリケーションガイド」が表示されている。なお、情報検索装置により、第1候補として、「マニュアル」が選択されている。

【0065】また、データベースの選択肢には、「マニュアルデータベース」、「入門書データベース」～が表示されている。また、第1候補として、処理部11により、「マニュアルデータベース」が選択されている。

【0066】検索者は、上記例のように表示されたキーワードの選択肢とデータベースの選択肢を選択して、検索条件を決定する。

【0067】図9の例では、図9(B)に示す検索条件で検索を行った結果、図9(C)に示すように、検索結果の表示がされている。この場合、該当件数が68件と多いいため、図9(C)に示すように、検索者は、検索条件の変更を選択して、検索範囲の絞り込みを選択し、絞り込みたい条件として、OSを選択している。

【0068】これにより、処理部11では、OSの元のキーワードである「UNIX」を下位語に展開して、下位概念のキーワードの一覧である、「UTS/M」、「UXP/M」、「SX/A」、「UXP/DS」、「SUN/OS」～を表示している。検索者は、この表示された下位概念のキーワードの一覧から、キーワードを選択する。これにより、検索対象を絞り込んで、再度、検索を行うことができる。

【0069】上記のように、第2実施例では、検索時に、検索者により入力された自然語の質問文から、キーワードの選択肢とデータベースの選択肢を生成し、検索者にこの選択肢を選択させて決定した検索条件で、文献データの検索を行う。このため、検索時に、SQL(Structured Query Language)等のような所定の質問式による記述が不要で、容易に適切なキーワードを生成させることができ、検索時のヒット率を向上させることができる。

【0070】

【発明の効果】上述の如く、請求項1の発明によれば、

文献データの登録時に、展開語辞書により自動的に適切な検索用キーワードを生成して登録するため、検索時の高いヒット率を実現しながら、検索時間を短縮することができ、また、検索用キーワードを付与する手間を削減することができる等の特長を有する。

【0071】請求項3の発明によれば、検索用キーワードによる文献データの検索の結果、該当する文献データの件数が基準値以上の場合には、下位概念の検索用キーワード候補を生成して検索者に選択させて、検索対象を絞り込むため、検索用キーワードに該当する文献データの件数が多い場合に、目的の文献データを検索するまでの検索時間を短縮することができる。

【0072】請求項4の発明では、検索時に、検索者により入力された自然語の質問文から、適切なキーワードを生成するため、容易に適切なキーワードを生成させることができ、検索時のヒット率を向上させることができる。

#### 【図面の簡単な説明】

【図1】本発明の第1実施例を適用した情報検索装置の構成図である。

【図2】第1実施例におけるキーワードの展開手順を示すフロー図である。

【図3】第1実施例におけるキーワード展開の一例の説明図である。

【図4】本発明の第1実施例における、検索手順を示すフロー図である。

【図5】第1実施例の検索時における下位概念のキーワードの生成例を示す図である。

\*

\* 【図6】本発明の第2実施例を適用した情報検索装置の構成図である。

【図7】本発明の第2実施例における、検索手順を示すフロー図である。

【図8】自然語の質問文の例と、検索条件確認時及び検索条件変更時の表示例を示す図である。

【図9】自然語の質問文の例と、検索条件確認時及び検索条件変更時の表示例を示す図である。

#### 【符号の説明】

- |    |                  |
|----|------------------|
| 10 | 1 1 処理部          |
|    | 1 2 文献データ入力装置    |
|    | 1 3 入力装置         |
|    | 1 4 表示装置         |
|    | 1 5 形態素辞書        |
|    | 1 6 不要語辞書        |
|    | 1 7 展開語辞書        |
|    | 1 8 上位語辞書        |
|    | 1 9 同義語辞書        |
|    | 2 0 複合語辞書        |
| 20 | 2 1 異表記語辞書       |
|    | 2 2 データベース定義情報   |
|    | 2 3 知識ベース        |
|    | 2 4 データベース群      |
|    | 2 5 データベース       |
|    | 2 5.~2 5. データベース |
|    | 2 6 文献データファイル    |
|    | 2 7 キーワードファイル    |

【図5】

第1実施例の検索時における  
下位概念のキーワードの生成例を示す図

	元のキーワード	生成されたキーワード
A	製造	化 学 機 構 ↓
B	保 险	生命保険 損害保険 ↓
C	パソコ	FMRシリーズ FM-TOWNS シリーズ ↓

【図8】

自然語の質問文の例と、検索条件確認時及び  
検索条件変更時の表示例を示す図

(A) 「食品販売店に食品を卸す食品卸業の物流管理と...  
在庫管理の事例と手法を探したい」

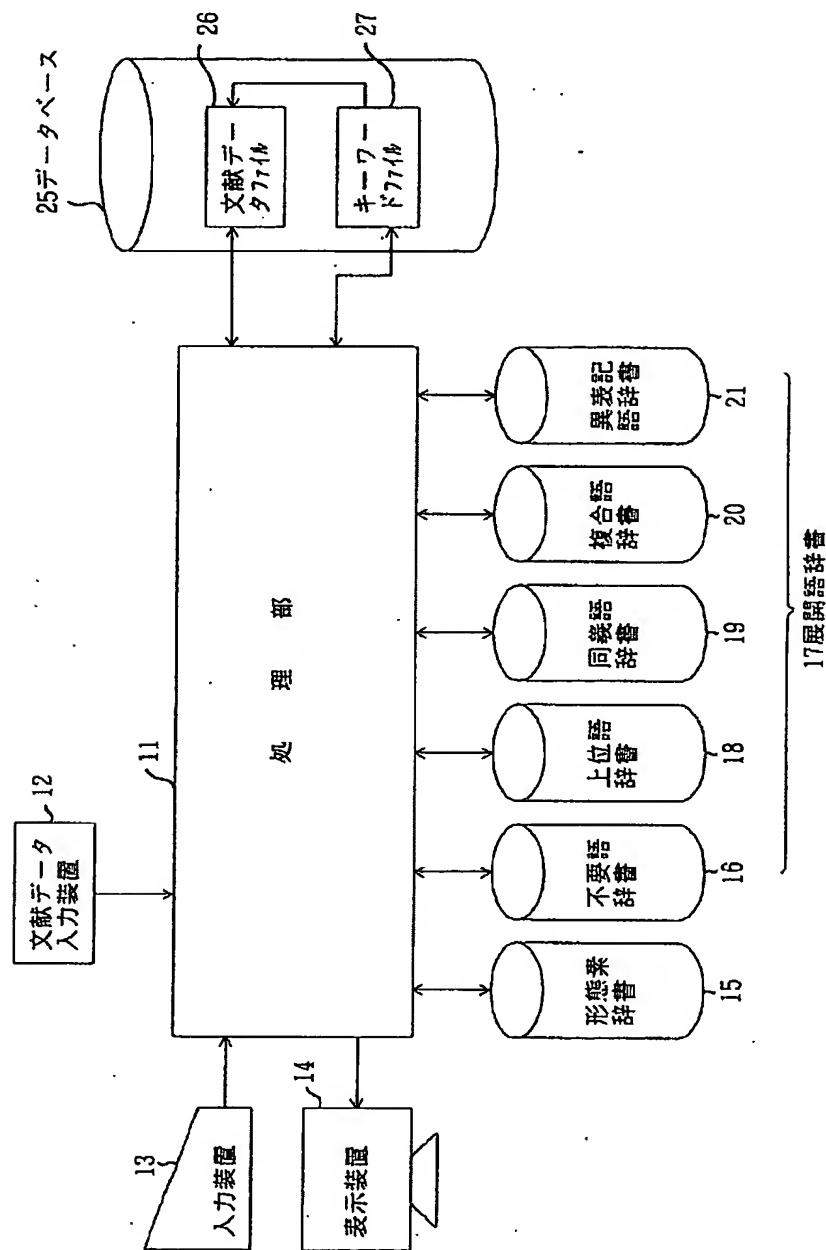
(B) 選択肢  
 検索キー :  食品卸  
 物流管理  
 在庫管理  
 手法  
 他の検索キー :  事例  
  
 データベース :  事例DB  
 提案書事例DB  
 請求書事例DB  
 ↓

(C) 1件見つかりました  
 検索条件を変更しますか?  
 YES  NO  
 戻り込み  検索範囲拡大  
 検索範囲を広げたい条件は?  
 品種  呼び名  他の検索キー  
 品種の一覧  
 流通  取扱元  
 ↓

→ 検索結果の表示  
 → 検索条件の変更を選択  
 → 検索範囲の拡大を選択  
 → 楽曲の範囲拡大を選択

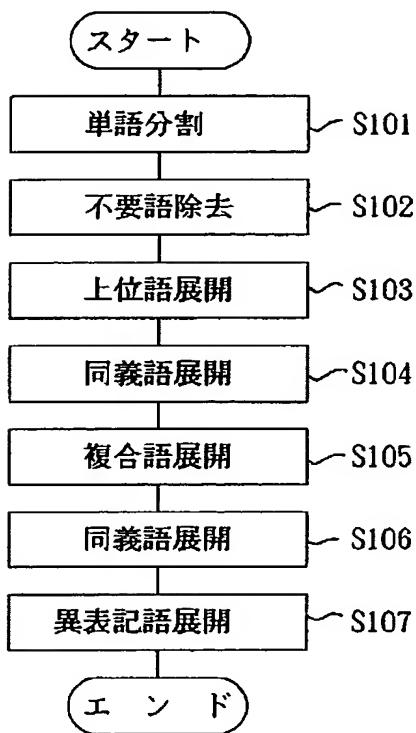
【図1】

## 本発明の第1実施例を適用した情報検索装置の構成図



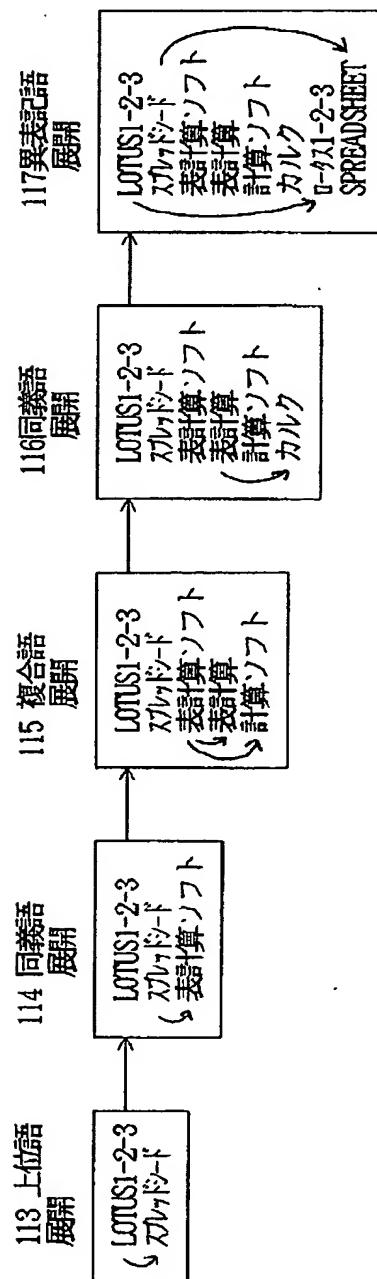
【図2】

第1実施例におけるキーワードの  
展開手順を示すフローチャート



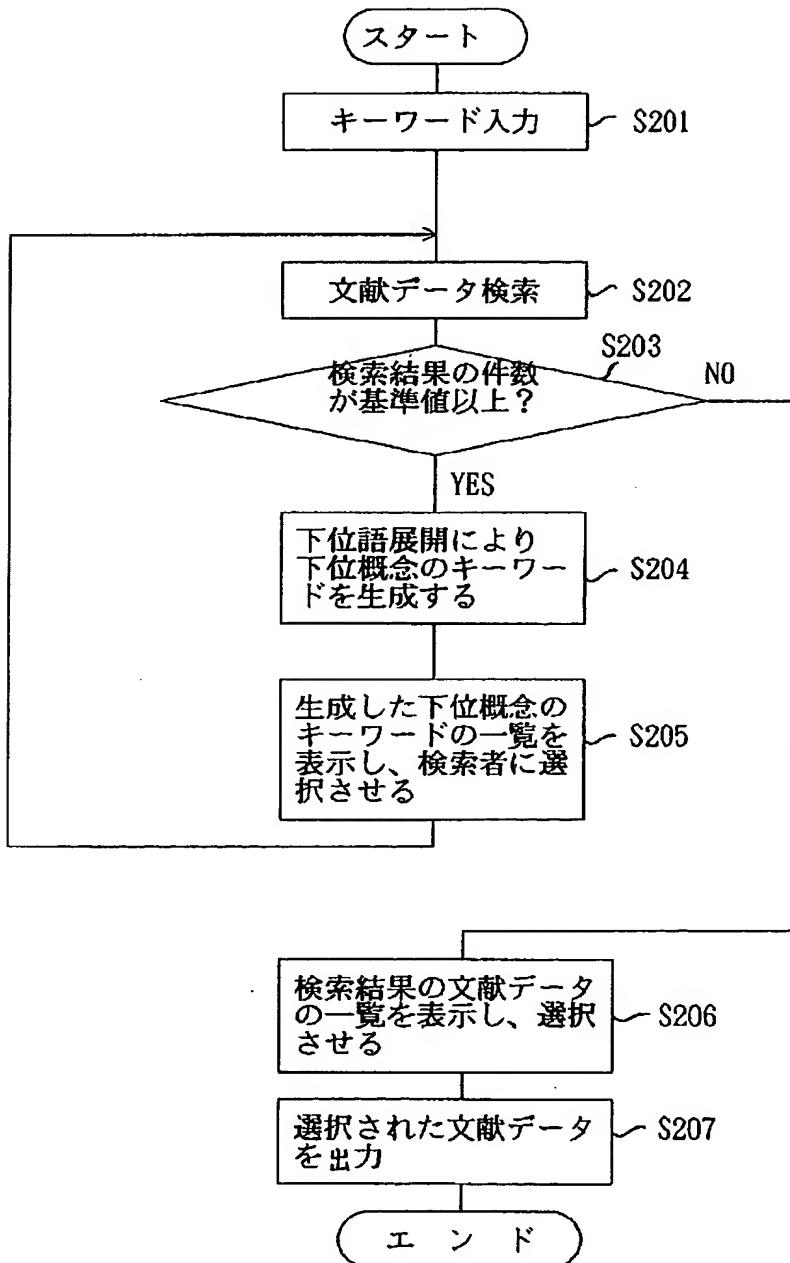
【図3】

## 第1実施例におけるキーワード展開の一例の説明図



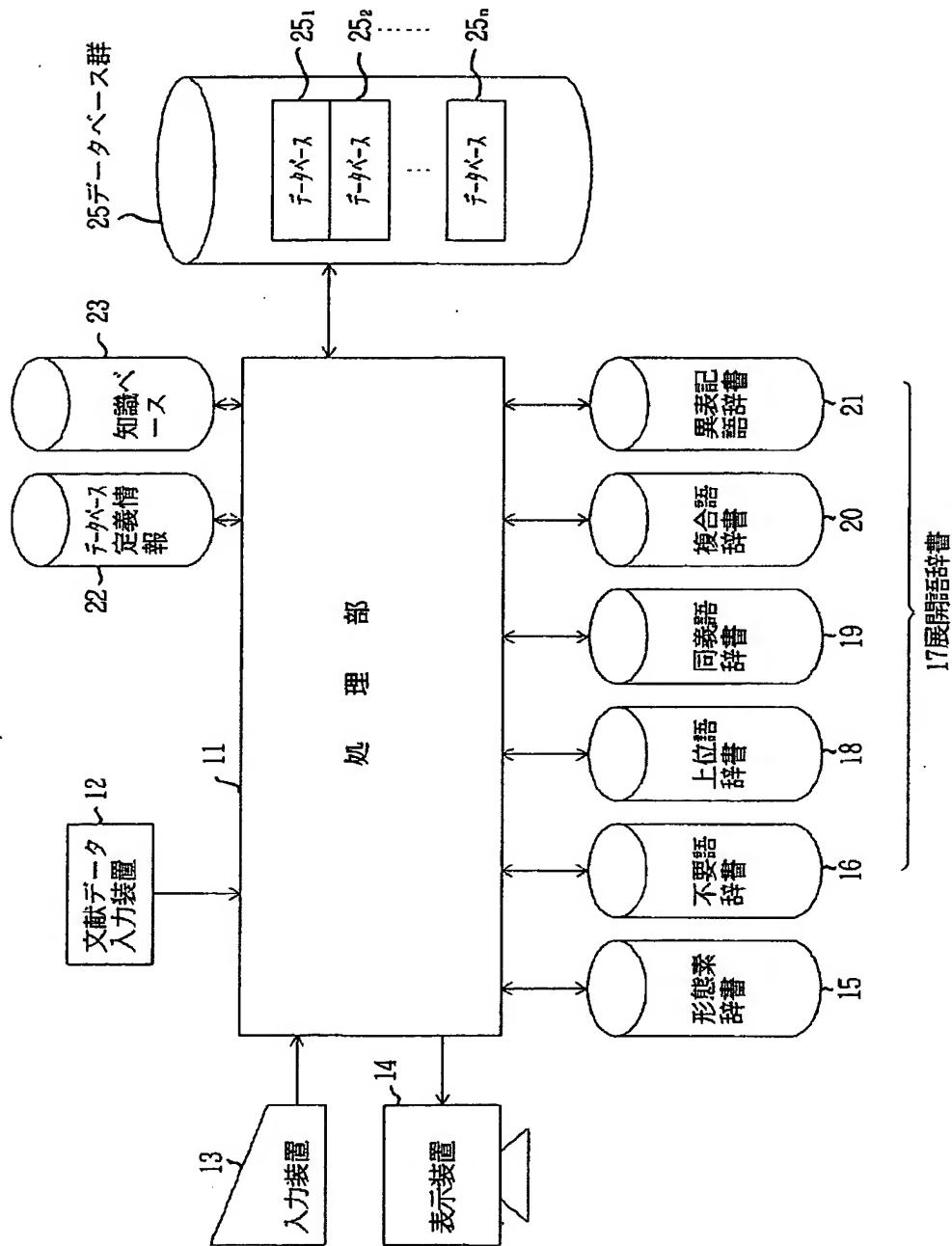
【図4】

本発明の第1実施例における  
検索手順を示すフローチャート



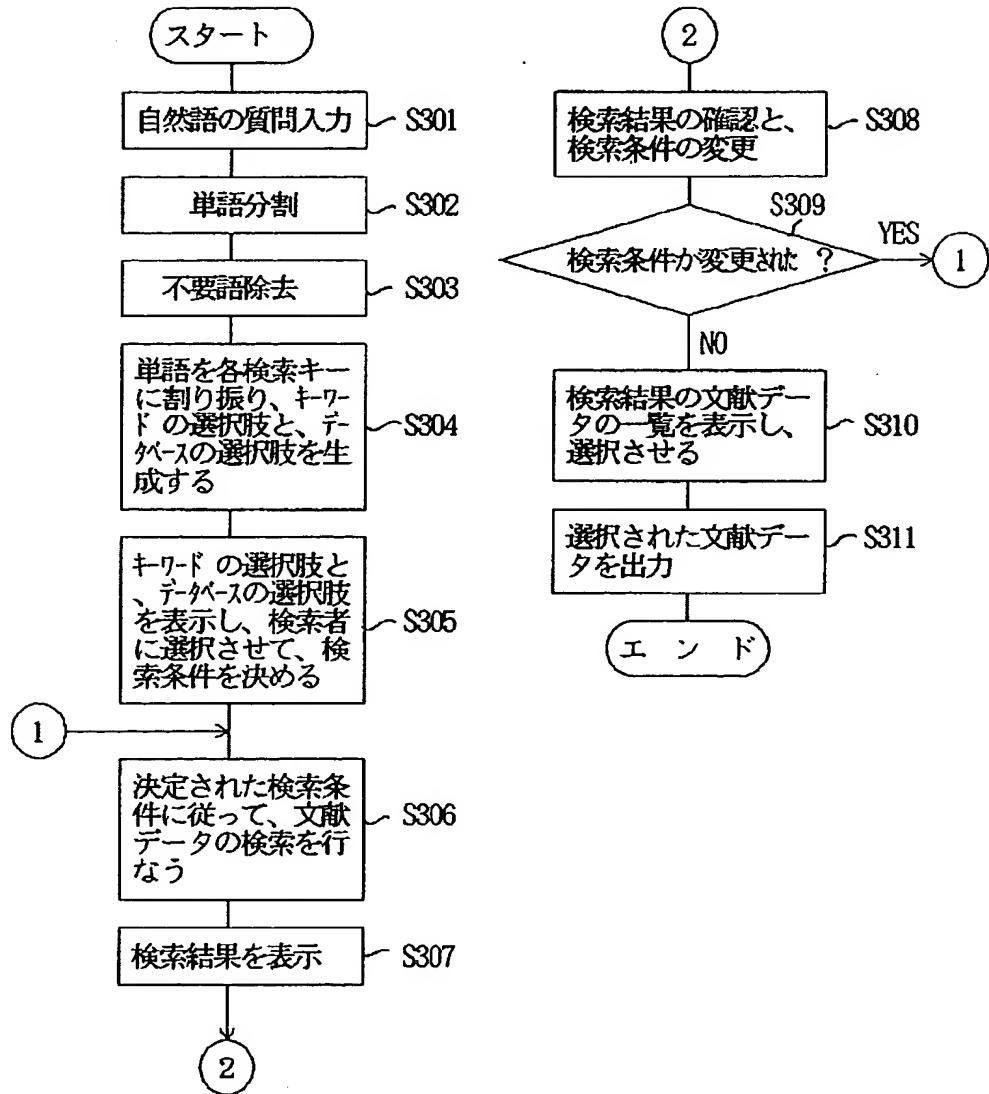
【図6】

## 本発明の第2実施例を適用した情報検索装置の構成図



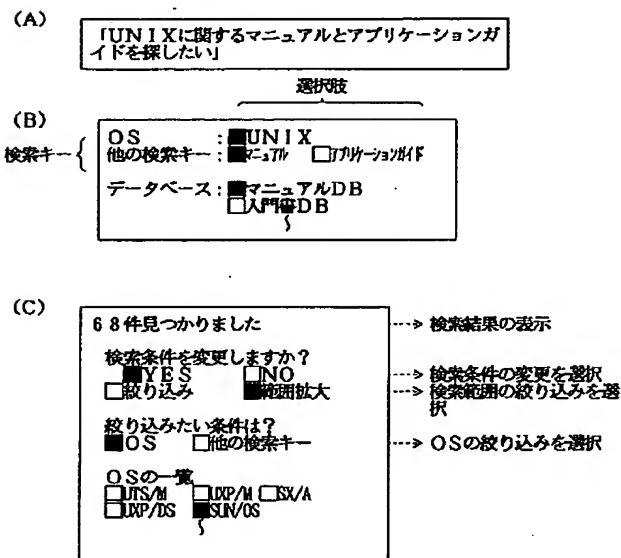
【図7】

本発明の第2実施例における  
検索の手順を示すフローチャート



【図9】

自然語の質問文の例と、検索条件固め時及び  
検索条件変更時の表示例を示す図



## フロントページの続き

(72)発明者 太田 竜夫

神奈川県川崎市中原区上小田中1015番地  
富士通株式会社内

(72)発明者 赤松 由理

神奈川県川崎市中原区上小田中1015番地  
富士通株式会社内